JOURNAL OF FUNCTIONAL ANALYSIS 39, 182–198 (1980)

Szegö's Limit Theorem: The Higher-Dimensional Matrix Case*

HAROLD WIDOM

University of California, Santa Cruz, California 95064

Communicated by the Editors

Received January 24, 1980; revised April 8, 1980

The continuous version of Szegö's theorem gives the first two terms of the asymptotics as $\alpha \to \infty$ of the determinants of certain convolution operators on $L_2(0, \alpha)$ with scalar-valued kernels. Generalizations are known if the kernel is matrix valued or if the interval $(0, \alpha)$ is replaced by $\alpha\Omega$ with Ω a bounded set in R^n with smooth boundary. This paper treats the higher-dimensional matrix case. The coefficient in the interesting (second) term is an integral over the contangent bundle of $\partial\Omega$ of the correponding coefficients of one-dimensional problems.

1. INTRODUCTION

The one-dimensional scalar-valued (continuous) version of Szegö's theorem [3] states that under appropriate conditions on the function $k(x)$ the operator T_α on $L_2(0, \alpha)$ defined by

$$T_\alpha f(x) = f(x) + \int_0^\alpha k(x-y) f(y)\, dy$$

has its determinant given asymptotically as $\alpha \to \infty$ by

$$\log \det T_\alpha = \alpha s(0) + \int_0^\infty z s(z)\, s(-z)\, dz + o(1), \tag{1}$$

where

$$s(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi z} \log \left\{ 1 + \int_{-\infty}^{\infty} e^{-i\xi x} k(x)\, dx \right\} d\xi.$$

In the higher-dimensional analogue the interval $[0, \alpha]$ is replaced by $\alpha\Omega$,

* Research supported by a grant from the National Science Foundation.

0022-1236/80/140182–17$02.00/0

where Ω is a compact subset of R^n with C^1 boundary. The result is, with obvious analogues of T_α, k, s,

$$\log\det T_\alpha = \alpha^n s(0)\,|\Omega| + \tfrac{1}{2}\alpha^{n-1}\int_{\partial\Omega} dx \int_{z\cdot\nu_x\geq 0} z\cdot\nu_x s(z)\,s(-z)\,dz + o(\alpha^{n-1}), \tag{2}$$

where $|\Omega|$ denotes the volume of Ω, ν_x is the inner unit normal to $\partial\Omega$ at the point x, and dx denotes surface measure. Three different proofs of this, under different hypotheses, may be found in [4, 5, 6].

In the one-dimensional matrix case $k(x)$ is matrix-valued and L_2 consists of vector-valued functions. If

$$\sigma(\xi) = 1 + \hat{k}(\xi) = 1 + \int_{-\infty}^{\infty} e^{-i\xi x}\,k(x)\,dx,$$

denote by $W(\sigma)$ the corresponding Wiener–Hopf operator on $L_2(R_+)$

$$W(\sigma)f(x) = f(x) + \int_0^\infty k(x-y)f(y)\,dy, \qquad x\in R_+.$$

Under appropriate conditions $\sigma(\xi)^{-1}$ will be of similar form, the product $W(\sigma)\,W(\sigma^{-1})$ will differ from I by a trace class operator (so its determinant will be defined), and the result is

$$\log\det T_\alpha = \frac{\alpha}{2\pi}\int_{-\infty}^{\infty}\log\det\sigma(\xi)\,d\xi + \log\det W(\sigma)\,W(\sigma^{-1}) + o(1). \tag{3}$$

The discrete (Toeplitz matrix) analogue of this was first proved in [8]; an elegant direct proof may be found in [1]. Both of these go over without difficulty to the continuous case.

Here we extend (3) to higher dimensions. It transpires that the symmetry of the appearance of σ and σ^{-1} in the last term of (3) is an accident of the symmetry of the underlying set, which for $n=1$ is an interval. An analogue of (3) cannot be expected to hold in general and we prove it in case Ω has a point of symmetry, or for any Ω if σ is even. In the generalization the expression

$$\log\det \mathrm{W}(\sigma)\,W(\sigma^{-1})$$

is replaced by an integral, over the cotangent bundle of $\partial\Omega$, of similar objects.

The cotangent bundle $\mathscr{X} = T^*(\partial\Omega)$ may be identified with the set of all pairs $X=(x, \eta)$, where $x \in \partial\Omega$ and $\eta \cdot \nu_x = 0$. There is a natural measure $dX = d\eta\, dx$ on $\mathscr{X}$, where $d\eta$ denotes Lebesgue measure on the hyperplane

$$\Pi_x = \{\eta : \eta \cdot \nu_x = 0\}$$

and dx is surface measure on $\partial\Omega$. Given $X=(x, \eta)$ we define

$$\sigma_X(\xi) = \sigma(\eta + \xi \nu_x), \qquad \xi \in R. \tag{4}$$

Then our formula reads

$$\log \det T_\alpha = \left(\frac{\alpha}{2\pi}\right)^n |\Omega| \int_{R^n} \log \det \sigma(\xi)\, d\xi + \frac{1}{2}\left(\frac{\alpha}{2\pi}\right)^{n-1} \int_{\mathscr{X}} \log \det W(\sigma_X)\, W(\sigma_X^{-1})\, dX + o(\alpha^{n-1}). \tag{5}$$

A formula which holds in general, i.e., without any symmetry condition, is

$$\log \det T_\alpha = \left(\frac{\alpha}{2\pi}\right)^n |\Omega| \int_{R^n} \log \det \sigma(\xi)\, d\xi + \left(\frac{\alpha}{2\pi}\right)^{n-1} \int_{\mathscr{X}} \operatorname{tr}\{\log W(\sigma_X) - W(\log \sigma_X)\}\, dX + o(\alpha^{n-1}). \tag{6}$$

This in turn is a consequence of a more general fact which is the main result of the paper. The basic assumptions are that

$$\|k\| \overset{\text{def}}{=} \int_{R^n} \|k(x)\|\, dx + \left\{\int_{R^n} |x|\, \|k(x)\|^2\, dx\right\}^{1/2} < \infty,$$

where $\|k(x)\|$ denotes the Hilbert–Schmidt norm of the matrix $k(x)$, and that

$$\hat{k}(\xi) \in L_1(R^n).$$

The last guarantees that $T_\alpha - I$ is of trace class.

THEOREM. *Suppose f is analytic on the closed convex hull Σ of the numerical ranges of the matrices $\sigma(\xi)$, $\xi \in R^n$, and $f(1) = 0$. Then*

$$\operatorname{tr} f(T_\alpha) = \left(\frac{\alpha}{2\pi}\right)^n |\Omega| \int_{R^n} \operatorname{tr} f(\sigma(\xi))\, d\xi + \left(\frac{\alpha}{2\pi}\right)^{n-1} \int_{\mathscr{X}} \operatorname{tr}\{f(W(\sigma_X)) - W(f(\sigma_X))\}\, dX + o(\alpha^{n-1}). \tag{7}$$

Recall that the numerical range of a Hilbert space operator A is the set of all (Au, u) with $\|u\| = 1$. The spectrum of A is cointained in the numerical range; in fact $\|(A - \lambda)^{-1}\|$ is at most the reciprocal of the distance from λ to the numerical range. We shall see that the assumption on f guarantees that every ingredient of (7) is well defined.

In case $f(w) = \log w$ then the hypothesis is satisfied if $0 \notin \Sigma$, for example, if $\|k\| < 1$, and (6) follows. The derivation of (5) from (6) under a symmetry condition, as well as the derivation of (2) from (6) in the scalar case, are matters of algebra—applications of the Baker–Campbell–Hausdorff formula. At the end of the paper we discuss relaxation of the condition $0 \notin \Sigma$.

The proof of the Theorem will run roughly as follows. We show that the inverses of operators like T_α or $W(\sigma_X)$ are, to a first approximation, given by replacing σ by σ^{-1}. This is an idea familiar to everyone. We deduce that for a family of σ depending on a complex parameter (specifically we replace σ by $1 - \lambda + \lambda\sigma$, where λ is in a complex neighborhood of $[0, 1]$) we have estimates

$$\operatorname{tr} f(T_\alpha) = \left(\frac{\alpha}{2\pi}\right)^n |\Omega| \cdot \int_{R^n} \operatorname{tr} f(\sigma(\xi))\, d\xi + O(\alpha^{n-1}), \tag{8}$$

$$\int_{\mathscr{X}} \operatorname{tr}\{f(W(\sigma_X)) - W(f(\sigma_X))\}\, dX = O(1), \tag{9}$$

which are uniform in λ. Thus by a normal families argument it will suffice to prove the result for small λ. For these we use the Taylor series for f, and the uniformity of (8) and (9) for certain sets of functions f, to reduce everything to the case where f is just a power. Here a computation is made, as in [3, 5]. The term that gives the second approximation is an integral over $\partial\Omega$, but rather than evaluating the integrand we interpret it as the integral over Π_x of the traces of operators on $L_2(R_+)$. This is the crux of the matter and constitutes the only new idea of the paper.

2. Proof of the Theorem

The matrix-valued distributions on R^n of the form

$$c\delta(x) + k(x)$$

with c a constant (times the identity matrix) and $\|k\| < \infty$ form a Banach algebra under convolution with the norm

$$|c| + \|k\|. \tag{10}$$

This is an easy exercise. Correspondingly the set $\hat{A}$ of Fourier transforms of elements of A

$$\sigma(\xi) = c + \hat{k}(\xi)$$

forms a Banach algebra under pointwise matrix multiplication with the same norm (10). Clearly an element of $\hat{A}$ is invertible exactly when its determinant is invertible in the corresponding scalar algebra, and this requires simply that the determinant be everywhere nonzero. Thus the spectrum of σ in $\hat{A}$ is the union of the spectra of the matrices $\sigma(\xi)$.

The Wiener–Hopf operator $W(\sigma)$ on $L_2(R_+)$ is the compression of convolution by $c\delta + k$ on $L_2(R)$. The convolution is unitarily equivalent to multiplication by σ. The numerical range of $W(\sigma)$ is therefore contained in the closed convex hull of the numerical ranges of the matrices $\sigma(\xi)$.

Another fact we shall use is that (with an inverted circumflex denoting the inverse Fourier transform) if

$$\int_{t \neq 0} |t| \, \|\check{\sigma}_i(t)\|^2 \, dt < \infty, \qquad i = 1, 2, \tag{11}$$

then $W(\sigma_1 \sigma_2) - W(\sigma_1) \, W(\sigma_2)$ is trace class and

$$\| W(\sigma_1 \sigma_2) - W(\sigma_1) \, W(\sigma_2) \|_1^2 \leqslant \int |t| \, \|\check{\sigma}_1(t)\|^2 \, dt \int |t| \, \|\check{\sigma}_2(t)\|^2 \, dt. \tag{12}$$

We deduce that if σ and σ^{-1} satisfy (11) and if $W(\sigma)$ is invertible then $W(\sigma)^{-1} - W(\sigma^{-1})$ is trace class and

$$\begin{aligned} &\| W(\sigma)^{-1} - W(\sigma^{-1}) \|_1^2 \\ &\qquad \leqslant \| W(\sigma)^{-1} \|^2 \int |t| \, \| \check{\sigma}(t) \|^2 \, dt \int |t| \, \| (\sigma^{-1})^{\vee} (t) \|^2 \, dt. \end{aligned} \tag{13}$$

If f is analytic on the closed convex hull of the numerical ranges of the $\sigma(\xi)$ then replacing σ by $\sigma - \lambda$, multiplying by $-f(\lambda)/2\pi i$, and integrating with respect to λ over a suitable contour show that $f(W(\sigma)) - W(f(\sigma))$ is trace class.

We return to the setting of the Theorem. The preceding discussion shows, first, that $f(\sigma(\xi))$ is well defined and $f(\sigma) \in \hat{A}$. Moreover, since $\sigma(\xi) \to 1$ as $|\xi| \to \infty$ and $f(1) = 0$ we also have

$$f(\sigma(\xi)) \in L_1(R^n).$$

Thus the first integral in (7) makes sense. We shall show next that

$$f(W(\sigma_X)) - W(f(\sigma_X))$$

is trace class for almost every $X \in \mathscr{X}$ and that its trace is integrable over $\mathscr{X}$. Let us compute, for $\sigma \in \hat{A}$ and $X = (x, \eta) \in \mathscr{X}$, the inverse Fourier transform $\check{\sigma}_X(t)$ of the function $\sigma_X(\xi)$ given by (4). We can write

$$\sigma_X(\xi) - c = \int_{R^n} e^{-i(\eta + \xi v_x) \cdot z} k(z)\, dz$$
$$= \int_{\Pi_x} e^{-i\eta \cdot y}\, dy \int_{-\infty}^{\infty} e^{-i\xi t} k(y + tv_x)\, dt$$

and so

$$\check{\sigma}_X(t) - c\delta(t) = \int_{\Pi_x} e^{-i\eta \cdot y} k(y + tv_x)\, dy. \tag{14}$$

Parseval's identity holds for matrix-valued functions if one uses their Hilbert–Schmidt norms and so

$$\int_{\Pi_x} \| \check{\sigma}_X(t) - c\delta(t)\|^2\, d\eta = (2\pi)^{n-1} \int_{\Pi_x} \| k(y + tv_x)\|^2\, dy.$$

Integrating with respect to t gives

$$\int_{\Pi_x} \int_{t \neq 0} |t|\, \| \check{\sigma}_X(t)\|^2\, dt\, d\eta$$
$$= (2\pi)^{n-1} \int_{-\infty}^{\infty} |t|\, dt \int_{\Pi_x} \| k(y + tv_x)\|^2\, dy$$
$$= (2\pi)^{n-1} \int_{R^n} |z \cdot v_x|\, \| k(z)\|^2\, dz.$$

The last integral is at most $\| k \|^2$ for all x. We deduce

$$\int_{\mathscr{X}} dX \int_{t \neq 0} |t|\, \| \check{\sigma}_X(t)\|^2\, dt \leqslant (2\pi)^{n-1} |\partial\Omega|\, \| \sigma \|^2. \tag{15}$$

We are now ready to apply (13). If we replace σ by $\sigma_X - \lambda$, integrate over $\mathscr{X}$, and use Schwarz's inequality, we find

$$\int_{\mathscr{X}} \| W(\sigma_X - \lambda)^{-1} - W(((\sigma_X - \lambda)^{-1}))\|_1\, dX$$
$$\leqslant (2\pi)^{n-1} |\partial\Omega| \sup_X \|(\sigma_X - \lambda)^{-1}\|\, \|\sigma - \lambda\|\, \|(\sigma - \lambda)^{-1}\|.$$

The spectrum of σ and the numerical ranges of the various $W(\sigma_X)$ are contained in Σ. Hence for λ on a curve surrounding Σ but still contained in the region of analyticity of f the right side is uniformly bounded. Multiplying by $f(\lambda)$ and integrating therefore give

$$\int_{\mathscr{X}} \|f(W(\sigma_X)) - W(f(\sigma_X))\|_1 \, dX < \infty. \tag{16}$$

Moreover, and this will be useful later, this holds uniformly for any set of functions f uniformly bounded and analytic in a convex open set U and any compact set of $\sigma \in \hat{A}$ such that all numerical ranges of the $\sigma(\xi)$ lie in U.

At this point we have not only shown that all ingredients of (7) make sense but we have also established (9). The proof of the preceding estimate (8) is quite similar. We begin it with a replacement for (12) for the operators T_α.

Given $\sigma \in \hat{A}$ denote by $C_\alpha(\sigma)$ convolution by $\check{\sigma}$ on $L_2(\alpha\Omega)$. Thus in this notation T_α is $C_\alpha(\sigma)$. If $\sigma_i = c_i + \hat{k}_i$ $(i = 1, 2)$ then the kernel of

$$C_\alpha(\sigma_1\sigma_2) - C_\alpha(\sigma_1)\, C_\alpha(\sigma_2)$$

at a point $(x, y) \in \alpha\Omega \times \alpha\Omega$ is equal to

$$\int_{\alpha\Omega^c} k_1(x-z)\, k_2(z-y)\, dz.$$

But

$$\int_{\alpha\Omega}\int_{\alpha\Omega^c} \|k_1(x-z)\|^2 \, dz\, dx = \int \|k_1(u)\|^2 \, |\alpha\Omega \cap (\alpha\Omega + u)^c| \, du.$$

The volume in the last integral is $O(\alpha^{n-1}|u|)$, the implied constant depending only on Ω. Hence

$$\int_{\alpha\Omega}\int_{\alpha\Omega^c} \|k_1(x-z)\|^2 \, dz\, dx = O(\alpha^{n-1} \|k_1\|^2).$$

A similar inequality holds with k_1 replaced by k_2 and so

$$\|C_\alpha(\sigma_1\sigma_2) - C_\alpha(\sigma_1)\, C_\alpha(\sigma_2)\|_1 = O(\alpha^{n-1} \|k_1\| \, \|k_2\|).$$

By the argument used above we deduce

$$\|f(C_\alpha(\sigma)) - C_\alpha(f(\sigma))\|_1 = O(\alpha^{n-1}),$$

and this holds uniformly for the same sets of f and σ as (16).

Note that $C_\alpha(f(\sigma))$ is itself trace class, with continuous kernel

$f(\sigma)^{\vee}(x-y)$. Its trace therefore equals the integral of the trace of its kernel over the diagonal in $\alpha\Omega \times \alpha\Omega$. Thus

$$\operatorname{tr} C_\alpha(f(\sigma)) = \left(\frac{\alpha}{2\pi}\right)^n |\Omega| \int \operatorname{tr} f(\sigma(\xi))\, d\xi. \tag{17}$$

This establishes (8) with uniformity as described after (16).

We proceed as indicated in the Introduction. Take the given f and σ and consider the family $\sigma_\lambda = 1 - \lambda + \lambda\sigma$, where λ runs over a small neighborhood of $[0, 1]$ in the complex plane. If the neighborhood is small enough the numerical ranges of all the $\sigma_\lambda(\xi)$ will lie in a fixed compact subset of the domain of f. Therefore (8) and (9) hold for σ_λ uniformly in λ. Since both integrals are analytic functions of λ the conclusion of the Theorem will follow forall λ, in particular $\lambda = 1$, if it can be proved for sufficiently small λ. In view of (17) what is to be proved is that for small λ

$$\begin{aligned}&\operatorname{tr}\{f(C_\alpha(\sigma)) - C_\alpha(f(\sigma))\}\\ &\quad = \left(\frac{\alpha}{2\pi}\right)^{n-1} \int_{\mathscr{X}} \operatorname{tr}\{f(W(\sigma_X)) - W(f(\sigma_X)\}\, dX + o(\alpha^{n-1})\end{aligned} \tag{18}$$

with σ now denoting σ_λ.

Let λ be so small that $1 - \lambda + \lambda\Sigma$ is contained in a fixed compact subset Σ' of the disc of convergence of the Taylor series for $f(w)$ about $w = 1$. Choose N so large that $R_N(w)$, the remainder after N terms of this series, satisfies

$$|R_N(w)| \leqslant \varepsilon$$

throughout a neighborhood of Σ'. It follows from (8) and (9) applied to the functions $\varepsilon^{-1}R_N(w)$ that

$$\operatorname{tr}\{R_N(C_\alpha(\sigma)) - C_\alpha(R_N(\sigma))\} = O(\varepsilon\alpha^{n-1})$$

$$\int_{\mathscr{X}} \operatorname{tr}\{R_N(W(\sigma_X)) - W(R_N(\sigma_X))\}\, dX = O(\varepsilon).$$

Since ε is arbitrarily small this implies that (18) holds for f if it holds for polynomials.

We shall now prove (18) for polynomials f. Replacing σ by $\sigma + c$ has the same effect as replacing w by $w + c$. Thus we may assume that $\sigma(\xi) = \hat{k}(\xi)$ and consider the case $f(w) = w^m$. We assume $m \geqslant 2$ since otherwise the result is trivial.

Denote by C convolution by k on $L_2(R^n)$ and for a given set $\Gamma \subset R^n$ let P denote multiplication by χ_Γ, the characteristic functions of Γ. (Thus

$C_\alpha(\sigma) = PCP$ when $\Gamma = \alpha\Omega$.) The kernel of $(PCP)^m$ at a point $(x, y) \in \Gamma \times \Gamma$ equals

$$\int \cdots \int k(x - x_1)\, k(x_1 - x_2) \cdots k(x_{m-2} - x_{m-1})$$
$$\times\, k(x_{m-1} - y)\, \chi_\Gamma(x_1) \cdots \chi_\Gamma(x_{m-1})\, dx_1 \cdots dx_{m-1},$$

the integration taken over $R^n \times \cdots \times R^n$. Change variables, replacing $x - x_1$, $x_1 - x_2, \ldots, x_{m-2} - x_{m-1}$ by $x_1, \ldots, x_{m-1}$, respectively. We obtain

$$\int \cdots \int k(x_1) \cdots k(x_m)\, k(x - y - x_1 - \cdots - x_{m-1})$$
$$\times\, \chi_\Gamma(x - x_1)\, \chi_\Gamma(x - x_1 - x_2) \cdots \chi_\Gamma(x - x_1 - \cdots - x_{m-1})\, dx_1 \cdots dx_{m-1}.$$

Since C^m has kernel $k * \cdots * k(x - y)$ we deduce that $PC^mP - (PCP)^m$ has kernel

$$\int \cdots \int k(x_1) \cdots k(x_m)\, k(x - y - x_1 - \cdots - x_{m-1})$$
$$\times\, \{1 - \chi_\Gamma(x - x_1) \cdots \chi_\Gamma(x - x_1 - \cdots - x_{m-1})\}\, dx_1 \cdots dx_{m-1}. \quad (19)$$

Take $\Gamma = \alpha\Omega$. Since $\sigma \in L_1$ the function k is continuous. It follows that the kernel is continuous and so to find the trace of the operator we integrate the trace of the kernel over the diagonal of $\alpha\Omega \times \alpha\Omega$. The result is

$$\operatorname{tr}\{C_\alpha(\sigma^m) - C_\alpha(\sigma)^m\}$$
$$= \int \cdots \int \operatorname{tr} k(x_1) \cdots k(x_m)\, k(-x_1 - \cdots - x_{m-1}) \quad (20)$$
$$\times\, |\alpha\Omega \backslash \alpha\Omega \cap (\alpha\Omega + x_1) \cap \cdots \cap (\alpha\Omega + x_1 + \cdots + x_{m-1})|\, dx_1 \cdots dx_{m-1}.$$

It is easy to show that for each i

$$\int \cdots \int \|k(x_1) \cdots k(x_{m-1})\, k(-x_1 - \cdots - x_{m-1})\|\, |x_i|\, dx_1 \cdots dx_{m-1}$$
$$\leqslant (m-1) \int |x|\, \|k(x)\|^2\, dx \left\{ \int \|k(x)\|\, dx \right\}^{m-2} \quad (21)$$

and so is finite. Moreover, the volume appearing in the integral in (20) is clearly

$$O(\alpha^{n-1} \max |x_i|),$$

and a little thought shows that it equals

$$\alpha^{n-1}\int_{\partial\Omega}\max(0, x_1\cdot\nu_x,..., (x_1+\cdots+x_{m-1})\cdot\nu_x)\,dx + o(\alpha^{n-1})$$

as $\alpha\to\infty$. This is a lemma left as an exercise for the reader. We therefore deduce from (20) that

$$\begin{aligned}&\operatorname{tr}\{C_\alpha(\sigma^m)-C_\alpha(\sigma)^m\}\\ &\quad=\alpha^{n-1}\int_{\partial\Omega}dx\int\cdots\int\operatorname{tr}k(x_1)\cdots k(x_{m-1})\,k(-x_1-\cdots-x_{m-1}) \qquad (22)\\ &\qquad\times\max(0, x_1\cdot\nu_x,..., (x_1+\cdots+x_{m-1})\cdot\nu_x)\,dx_1\cdots dx_{m-1}+o(\alpha^{n-1}).\end{aligned}$$

We shall now show how this integral may be rewritten as an integral over $\mathscr{X}$. In the inner integral over $R^n\times\cdots\times R^n$ write

$$x_i=y_i+t_i\nu_x,\qquad t_i\in R,\quad y_i\in\Pi_x.$$

Observe that

$$\begin{aligned}&\max(0, x_1\cdot\nu_x,..., (x_1+\cdots+x_{m-1})\cdot\nu_x)\\ &\quad=\max(0, t_1,..., t_1+\cdots+t_{m-1})\end{aligned}$$

and so in particular is independent of the y_i, with respect to which we shall integrate first. The integral

$$\begin{aligned}&\int\cdots\int k(y_1+t_1\nu_x)\cdots k(y_{m-1}+t_{m-1}\nu_x)\\ &\qquad\times k(-y_1-t_1\nu_x-\cdots-y_{m-1}-t_{m-1}\nu_x)\,dy_1\cdots dy_{m-1} \qquad (23)\end{aligned}$$

is the value at 0 of the convolution of the functions of y

$$k(y+t_i\nu_x),\qquad i\leqslant m-1,\qquad k(y-t_1\nu_x-\cdots-t_{m-1}\nu_x).$$

The Fourier transform of $k(y+t\nu_x)$ equals $\check{\sigma}_X(t)$ thought of as a function of $\eta\in\Pi_x$. (Recall that $X=(x,\eta)$.) It follows from the assumptions $k\in L_1$, $\sigma\in L_1$ that for almost every t

$$k(y+t\nu_x),\qquad \check{\sigma}_X(t),$$

belong to L_1 as functions of y and η, respectively. This is enough to guarantee that the convolution (23) equals

$$\frac{1}{(2\pi)^{n-1}}\int_{\Pi_x}\check{\sigma}_X(t_1)\cdots\check{\sigma}_X(t_{m-1})\,\check{\sigma}_X(-t_1-\cdots-t_{m-1})\,d\eta$$

for almost every $t_1,..., t_{m-1}$.

This is then integrated with respect to $t_1,\dots,t_{m-1}$ over $R\times\cdots\times R$ and then with respect to x over $\partial\Omega$. We conclude that (22) may be rewritten

$$\begin{aligned}&\operatorname{tr}\{C_\alpha(\sigma^m)-C_\alpha(\sigma)^m\}\\&\quad=\left(\frac{\alpha}{2\pi}\right)^{n-1}\int_{\mathscr{X}} dX\int\cdots\int \operatorname{tr}\check{\sigma}_X(t_1)\cdots\check{\sigma}_X(t_{m-1})\,\check{\sigma}_X(-t_1-\cdots-t_{m-1})\\&\qquad\times\max(0,t_1,\dots,t_1+\cdots+t_{m-1})\,dt_1\cdots dt_{m-1}+o(\alpha^{n-1}).\end{aligned}\tag{24}$$

This involved an interchange of order of integration which is justified by Fubini's theorem since for each i

$$\begin{aligned}&\int\cdots\int\|\check{\sigma}_X(t_1)\cdots\check{\sigma}_X(t_{m-1})\,\check{\sigma}_X(-t_1-\cdots-t_{m-1})\|\,|t_i|\,dt_1\cdots dt_{m-1}\\&\qquad\leqslant(m-1)\int|t|\,\|\check{\sigma}_X(t)\|^2\,dt\left\{\int\|\check{\sigma}_X(t)\|\,dt\right\}^{m-2}\end{aligned}$$

by (21), since for each X

$$\int_{-\infty}^{\infty}\|\check{\sigma}_X(t)\|\,dt\leqslant\int_{R^n}\|k(z)\|\,dz$$

by (14), and since

$$\int_{\mathscr{X}} dX\int|t|\,\|\check{\sigma}_X(t)\|^2\,dt<\infty$$

by (15).

Finally we compute the trace of $W(\sigma_X^m)-W(\sigma_X)^m$. The last displayed formulas imply that $\sigma_X(t)$ belongs to the one-dimensional algebra A for almost every X. Therefore the kernel of $W(\sigma_X^m)-W(\sigma_X)^m$ is given by an analogue of (19); the variables $x_i\in R^n$ are replaced by $t_i\in R$, the function $k(x)$ is replaced by $\sigma_X(t)$, and the set Γ by R_+. For almost every X the function $\sigma_X(\xi)$ belongs to L_1 and for these X the kernel is continuous. So to find the trace of the operator we integrate the trace of the kernel over the diagonal of $R_+\times R_+$ and find

$$\begin{aligned}&\operatorname{tr}\{W(\sigma_X^m)-W(\sigma_X)^m\}\\&\quad=\int\cdots\int\operatorname{tr}\check{\sigma}_X(t_1)\cdots\check{\sigma}_X(t_{m-1})\,\check{\sigma}_X(-t_1-\cdots-t_{m-1})\\&\qquad\times\max(0,t_1,\dots,t_1+\cdots+t_{m-1})\,dt_1\cdots dt_{m-1}.\end{aligned}$$

Comparing this with (24) completes the proof of the Theorem.

3. Derivation of (2) and (5)

The Baker–Campbell–Hausdorff formula [2, Chap. X] says that if α and β are elements of a Banach algebra of sufficiently small norm then

$$\log e^{\alpha} e^{\beta} = \sum_{k=1}^{\infty} u_k(\alpha, \beta),$$

where each $u_k(\alpha, \beta)$ is a linear combination of $(k-1)$-fold commutators

$$[\cdots [[\gamma_1 \gamma_2] \gamma_3] \cdots \gamma_k]$$

with each γ_i equal to α or β. The first few terms of the series are

$$\alpha + \beta + \frac{1}{2}[\alpha\beta] + \frac{1}{12}[[\alpha\beta]\beta] + \frac{1}{12}[[\beta\alpha]\alpha] + \cdots.$$

Let σ be an element of the one-dimensional algebra $\hat{A}$ satisfying $\|1-\sigma\| < 1$, so that $\log \sigma$ is well defined. We shall show that if $\|\log \sigma\|$ is sufficiently small (in other words, if $\|1-\sigma\|$ is sufficiently small) there are elements φ_+, φ_- of the algebras $\hat{A}_+, \hat{A}_-$, where

$$A_+ = \{c\delta + k\colon k(x) = 0 \text{ for } x < 0\},$$
$$A_- = \{c\delta + k\colon k(x) = 0 \text{ for } x > 0\},$$

such that

$$\sigma = e^{\varphi_-} e^{\varphi_+}, \tag{25}$$

or equivalently, such that

$$\log \sigma = \sum_{k=1}^{\infty} u_k(\varphi_-, \varphi_+). \tag{26}$$

Denote by $\varphi \to \varphi_+$ the obvious projection from $\hat{A}$ onto $\hat{A}_+$, set $\varphi_- = \varphi - \varphi_+$, and define

$$U(\varphi) = \log \sigma - \sum_{k=2}^{\infty} \lambda^{k-1} u_k(\varphi_-, \varphi_+).$$

Then $\lim_{N\to\infty} U^N(0)$ has a formal power series expansion

$$\log \sigma - \frac{\lambda}{2}[(\log \sigma)_-, (\log \sigma)_+] + \cdots,$$

which converges to an element of $\hat{A}$ satisfying $U(\varphi) = \varphi$ if λ is small enough, or for $\lambda = 1$ if $\|\log \sigma\|$ is small enough. Assuming the latter we set $\lambda = 1$ and find that

$$\log \sigma = \varphi + \sum_{k=2}^{\infty} u_k(\varphi_-, \varphi_+) = \sum_{k=1}^{\infty} u_k(\varphi_-, \varphi_+)$$

since $\varphi = \varphi_- + \varphi_+$. This gives (25).

If $\sigma_1 \in \hat{A}_-$ or $\sigma_2 \in \hat{A}_+$ then $W(\sigma_1 \sigma_2) = W(\sigma_1)\, W(\sigma_2)$. It follows from this fact and (25) that

$$W(\sigma) = e^{W(\varphi_-)}\, e^{W(\varphi_+)}$$

and so, again if $\|\log \sigma\|$ is sufficiently small,

$$\log W(\sigma) = \sum_{k=1}^{\infty} u_k(W(\varphi_-), W(\varphi_+)).$$

Using (26) we deduce

$$\begin{aligned} &\log W(\sigma) - W(\log \sigma) \\ &\quad = \sum_{k=2}^{\infty} \{u_k(W(\varphi_-), W(\varphi_+)) - W(u_k(\varphi_-, \varphi_+))\}, \end{aligned} \tag{27}$$

the terms corresponding to $k = 1$ cancelling.

We now return to R^n. By the usual argument we need to prove (2) or (5) for the family

$$\sigma_\lambda = 1 - \lambda + \lambda \sigma$$

for sufficiently small λ. If λ is small enough then for almost every X (27) will hold with σ replaced by $(\sigma_\lambda)_X$. So we assume λ is this small, suppress its appearance, and consider the second integral in (6).

In the scalar case all $u_k(\varphi_-, \varphi_+)$ for $k \geqslant 2$ vanish. Moreover, the commutator of Wiener–Hopf operators is trace class in this case and so all higher commutators have trace zero. Therefore

$$\begin{aligned} &\operatorname{tr}\{\log W(\varphi_X) - W(\log \sigma_X)\} \\ &\quad = \operatorname{tr} u_2(W(\varphi_-), W(\varphi_+)) = \tfrac{1}{2}\operatorname{tr}\{W(\varphi_+\varphi_-) - W(\varphi_+)\, W(\varphi_-)\} \\ &\quad = \tfrac{1}{2}\int_0^{\infty} t\check{\varphi}_+(t)\, \check{\varphi}_-(-t)\, dt = \tfrac{1}{2}\int_0^{\infty} t\check{\varphi}(t)\, \check{\varphi}(-t)\, dt. \end{aligned}$$

Since $\check{\varphi} = (\log \sigma_X)^\vee$ integration with respect to $\eta \in \Pi_x$ gives, as in the derivation of (24) from (22),

$$\frac{1}{(2\pi)^{n-1}} \int_{\Pi_x} \operatorname{tr}\{\log W(\sigma_X) - W(\log \sigma_X)\}\, d\eta$$

$$= \tfrac{1}{2} \int_{z \cdot v_x > 0} z \cdot v_x s(z)\, s(-z)\, dz$$

and so (2) follows from (6).

Now for the derivation of (5) under a symmetry condition. We return to one dimension and formula (27), which we shall write in an equivalent form. Define the function $\tilde{\sigma}$ by $\tilde{\sigma}(\xi) = \sigma(-\xi)$. From (25) we obtain

$$\tilde{\sigma}^{-1} = e^{-\tilde{\varphi}_+} e^{-\tilde{\varphi}_-}$$

and $\tilde{\varphi}_+, \tilde{\varphi}_-$ belong to the algebras $\hat{A}_-, \hat{A}_+$. Hence applying (27) to $\tilde{\sigma}^{-1}$ gives

$$\begin{aligned}
&\log W(\tilde{\sigma}^{-1}) + W(\log \tilde{\sigma}) \\
&\quad = \sum_{k=2}^{\infty} \{u_k(W(-\tilde{\varphi}_+), W(-\tilde{\varphi}_-)) - W(u_k(-\tilde{\varphi}_+, -\tilde{\varphi}_-))\} \\
&\quad = -\sum_{k=1}^{\infty} \{u_k(W(\tilde{\varphi}_-), W(\tilde{\varphi}_+)) - W(u_k(\tilde{\varphi}_-, \tilde{\varphi}_+))\}.
\end{aligned} \tag{28}$$

Here we have used the fact $u_k(-\alpha, -\beta) = -u_k(\beta, \alpha)$, a reflection of

$$\log e^{-\alpha} e^{-\beta} = -\log e^{\beta} e^{\alpha}.$$

Now for any $\varphi, \psi \in A$ the operator $W(\varphi) W(\psi) - W(\varphi\psi)$ is trace class and

$$\operatorname{tr}\{W(\varphi) W(\psi) - W(\varphi\psi)\} = -\int_0^{\infty} z \operatorname{tr} \check{\varphi}(z)\, \check{\psi}(-z)\, dz.$$

Consequently

$$\operatorname{tr}\{[W(\varphi), W(\psi)] - W([\varphi, \psi])\} = -\int_{-\infty}^{\infty} z \operatorname{tr} \check{\varphi}(z)\, \check{\psi}(-z)\, dz.$$

Replacing φ, ψ by $\tilde{\varphi}$, $\tilde{\psi}$ changes the sign of the right side. Consequently

$$\begin{aligned}
&\operatorname{tr}\{[W(\tilde{\varphi}), W(\tilde{\psi})] - W([\tilde{\varphi}, \tilde{\psi}])\} \\
&\quad = -\operatorname{tr}\{[W(\varphi), W(\psi)] - W([\varphi, \psi])\}.
\end{aligned}$$

Induction shows that this extends to higher commutators, and so

$$\begin{aligned}&\operatorname{tr}\{u_k(W(\tilde\varphi), W(\tilde\psi)) - W(u_k(\tilde\varphi, \tilde\psi))\}\\ &\qquad = -\operatorname{tr}\{u_k(W(\varphi), W(\psi)) - W(u_k(\varphi, \psi))\}.\end{aligned}$$

Therefore going back to (28) and comparing with (27) show that

$$\operatorname{tr}\{\log W(\sigma) - W(\log\sigma)\} = \operatorname{tr}\{\log W(\tilde\sigma^{-1}) + W(\log\tilde\sigma)\}$$

and so also

$$\begin{aligned}&\operatorname{tr}\{\log W(\sigma) - W(\log\sigma)\}\\ &\quad = \tfrac{1}{2}\operatorname{tr}\{\log W(\sigma) - W(\log\sigma) + \log W(\tilde\sigma^{-1}) + W(\log\tilde\sigma)\}.\end{aligned} \tag{29}$$

Now back to the second integral in (6). We must replace each σ above by σ_X and integrate with respect to X. The contributions of the first two terms on the right side of (29) we leave unchanged as

$$\tfrac{1}{2}\int_{\mathscr{X}} \operatorname{tr}\{\log W(\sigma_X) - W(\log\sigma_X)\}\, dX.$$

As for the last two terms, a simple check shows that if $\tilde X = (x, -\eta)$ then

$$(\sigma_X)^{\sim} = \tilde\sigma_{\tilde X}.$$

Therefore the variable change $X \to \tilde X$, under which the measure dX is invariant, gives for the contribution of the last two terms on the right side of (29)

$$\tfrac{1}{2}\int_{\mathscr{X}} \operatorname{tr}\{\log W(\tilde\sigma_X^{-1}) + W(\log\tilde\sigma_X)\}\, dX.$$

Therefore the second integral in (6) is equal to

$$\begin{aligned}\tfrac{1}{2}\int_{\mathscr{X}} \operatorname{tr}\{\log W(\sigma_X) + \log W(\tilde\sigma_X^{-1}) - W(\log\sigma_X)&\\ + W(\log\tilde\sigma_X)\}\, dX.&\end{aligned}$$

If σ is even then of course $\tilde\sigma = \sigma$ and the above expression equals

$$\tfrac{1}{2}\int_{\mathscr{X}} \operatorname{tr}\{\log W(\sigma_X) + \log W(\sigma_X^{-1})\}\, dX$$

and this gives (5).

If Ω is symmetric about the origin and if we write $-X=(-x,-\eta)$ then

$$\tilde{\sigma}_X = \sigma_{-X}$$

since $\nu_{-x} = -\nu_x$. Thus the change of variable $X \to -X$ gives

$$\int_{\mathscr{X}} \operatorname{tr}\{\log W(\tilde{\sigma}_X^{-1}) + W(\log \tilde{\sigma}_X)\}\, dX$$

$$= \int_{\mathscr{X}} \operatorname{tr}\{\log W(\sigma_X^{-1}) + W(\log \sigma_X)\}\, dX$$

and again (5) follows. The translation-invariance of the situation shows that (5) holds whatever point of symmetry Ω may have.

4. Concluding Remarks

What of (5) or (6) if the condition $0 \notin \Sigma$ is not satisfied? The situation is this. Let

$$\mathscr{A} = \{\sigma \in \hat{A} \colon W(\sigma_X) \text{ is invertible for all } X\}.$$

This is an open subset of $\hat{A}$. If $n=1$ then $\mathscr{A}$ is connected (the discrete analogue of this was proved in [7]) and may be so in general. In any case it is almost certainly true that if σ is in the connected component of $\mathscr{A}$ containing I then (5) under a symmetry condition, and (6) suitably interpreted, hold.

Here is why. Let σ_λ be a piecewise analytic curve in $\mathscr{A}$ with $\sigma_0 = 1$ and $\sigma_1 = \sigma$. If one defines

$$\log \sigma_\lambda = \int_0^\lambda \sigma_r' \sigma_r^{-1}\, dr,$$

$$\log \mathrm{W}((\sigma_\lambda)_X) = \int_0^\lambda W((\sigma_r')_X)\, W((\sigma_r)_X)^{-1}\, dr$$

then these are not logarithms in the usual sense, even when logarithms exist, and they also depend on the particular curve chosen. Nevertheless if λ is sufficiently small

$$\operatorname{tr}\{\log W((\sigma_\lambda)_X) - W(\log(\sigma_\lambda)_X)\} \tag{30}$$

is the same whether the above logarithms or the actual ones are used. The second integrand in (6) is to be interpreted as (30) with $\lambda = 1$.

As for the second integrand in (5) there is no difficulty about its meaning, for

$$\det W(\sigma_X)\, W(\sigma_X^{-1})$$

has limit 1 as $|\eta| \to \infty$ and the one-point compactification of $\mathscr{X}$ is simply connected. Hence there is a unique continuous logarithm with limit zero at infinity and this is the integrand.

Since (5) and (6) are true for λ small what one needs to prove them in general are estimates of the form (8) and (9) valid for λ in a neighborhood of $[0, 1]$. If one recalls how these were obtained earlier one sees that there is no difficulty with (9) but that (8) requires knowing that if $\sigma \in \mathcal{O}\!\!\!\mathcal{I}$ then the (operator) norms of the T_α^{-1} are bounded as $\alpha \to \infty$. In the scalar case this was proved in [6]. The argument given there goes over to the matrix case except for one awkward point. One has to know that the norms of the kernels of the $W(\sigma_X)^{-1}$ are bounded by

$$c\delta(x-y) + j(x-y)$$

for some $c\delta + j$ belonging to scalar-valued A and independent of X. In the scalar case this was automatic and probably is here, too, but we have not been able to prove it without an additional assumption. One that will serve is that the function

$$k^*(x) = \sup\{\|k(y)\| : |y| = |x|\}$$

belongs to scalar-valued A. Thus if $\sigma \in \mathcal{O}\!\!\!\mathcal{I}$ and satisfies this condition then (5) holds if there is symmetry and (6) if suitably interpreted.

References

1. E. Basor and J. W. Helton, A new proof of the Szegö limit theorem and new results for Toeplitz operators with discontinuous symbol, *J. Operator Theory* **3** (1980), 23–40.
2. G. Hochschild, "The Structure of Lie Groups," Holden–Day, San Francisco, 1965.
3. M. Kac, Toeplitz matrices, translation kernels, and a related problem in probability theory. *Duke Math. J.* **21** (1954), 501–509.
4. I. Ju. Linnik, A multidimensional analogue of a limit theorem of G. Szegö, *Math. USSR-Izv.* **9** (1975), 1323–1332.
5. H. Widom, A theorem on translation kernels in *n* dimensions, *Trans. Amer. Math. Soc.* **94** (1960), 170–180.
6. H. Widom, Asymptotic inversion of convolution operators, *Publ. Math. I.H.E.S.* **44** (1975), 191–240.
7. H. Widom, Asymptotic behavior of block Toeplitz matrices and determinants, *Advances in Math.* **13** (1974), 284–322.
8. H. Widom, On the limit of block Toeplitz determinants, *Proc. Amer. Math. Soc.* **50** (1975), 167–173.